PRB

KIDS COUNT
Data Resource Guide

By Alicia VanOrman

Table of Contents

General Resources

Abbreviations

ACS: American Community Survey

ACS PUMS: American Community Survey Public Use Microdata Sample

API: application programming interface

CDC: Centers for Disease Control and Prevention

CI: confidence interval

CPS: Current Population Survey

CV: coefficient of variation

MDAT: Microdata Access Tool

MOE: margin of error

NSCH: National Survey of Children's Health

PEP: Population Estimates Program

PUMA: Public Use Microdata Area

SAHIE: Small Area Health Insurance Estimates program

SAIPE: Small Area Income and Poverty Estimates program

SE: standard error

NOTE THANUN, VIA UNSPLASH

Introduction

KIDS COUNT works to improve the lives of children throughout the United States by using reliable evidence and analysis to advance sound policy reforms. KIDS COUNT does this, in part, by providing the best available data on children and their families at the national, state, and local levels through the KIDS COUNT Data Center, the *KIDS COUNT Data Book*, special reports, and blog posts.

This *KIDS COUNT Data Resource Guide* provides the KIDS COUNT network with a snapshot of the data sources used by the national KIDS COUNT program as well as some best practices for using them. These sources provide up-to-date and reliable information that is comparable across all 50 states, the District of Columbia, and, where available, Puerto Rico and the U.S. Virgin Islands. Together, they inform over 300 indicators on the KIDS COUNT Data Center and provide vital inputs for the *KIDS COUNT Data Book* and other KIDS COUNT publications.

KIDS COUNT state organizations and partners use some of the sources featured in this guide to gather data on the well-being of children and families in their local communities. And many of the tools and best practices discussed herein will likely apply across locations and communities.

This guide is designed to be used electronically and interactively, but may also be used as a standalone resource. It is organized into four sections. **Section one** provides a brief list of links to key data sources, related guidelines and resources, and tools to support data collection and analysis. It's a good place to start for users who know what data they're looking for but aren't sure where to find them.

Section two gives an overview of each of the major data sources used by the KIDS COUNT program, including for the KIDS COUNT Data Center, the *KIDS COUNT Data Book*, and special reports. The discussion describes the data that are available in each source, how to access that data, and some best practices for using each source. This section is designed to help data users get started with the sources and to support further learning.

Section three provides a list of the 16 key indicators used in the *KIDS COUNT Data Book* and the corresponding data sources.

Section four focuses on broad best practices that apply across data sources. Topics include assessing the reliability of estimates, comparing estimates, and disaggregating data by race and ethnicity. Rather than provide comprehensive best practices, this section offers highlights and information on where to learn more.

Quick Links to Data Sources and Tools

SOURCES AND GUIDANCE

- KIDS COUNT Data Center
- American Community Survey
- Population and Housing Unit Estimates (Population Estimates Program)
- Decennial Census
- Current Population Survey
- National Survey of Children's Health
- National Center for Education Statistics
- National Vital Statistics System

TOOLS AND GUIDANCE

U.S. CENSUS BUREAU

- Data.census.gov
- Microdata Access Tool
- Available APIs
- Census Academy (webinars and training)
- Comparing ACS Data
- Statistical Testing Tool

IPUMS

- IPUMS
- IPUMS USA
- IPUMS CPS
- IPUMS NHGIS

GEOGRAPHY INFORMATION AND TOOLS

- Guidance for Geography Users (U.S. Census Bureau)
- TIGERweb (U.S. Census Bureau)
- Geocorr (Missouri Census Data Center)

CENTERS FOR DISEASE CONTROL AND PREVENTION

- CDC WONDER
- CDC WISQARS

R

- R for Data Science (Hadley Wickham and Garrett Grolemund)
- Analyzing US Census Data: Methods, Maps, and Models in R (Kyle Walker)
- R (resource list from UCLA's Statistical Consulting Group)

National KIDS COUNT Data Sources

The national KIDS COUNT program relies on data from a variety of sources, most of them within the federal statistical system. These sources must be reliable, updated regularly, and comparable across all states.

Decennial Census

Conducted by the U.S. Census Bureau, the decennial census is a **once-a-decade effort to count the residents of the United States** and its five territories. The decennial census is an especially important dataset because it is used in ways that impact daily life in the United States, including:

- To determine apportionment of the U.S. House of Representatives.
- To inform congressional and legislative redistricting.
- To allocate more than **$1 trillion** in annual federal funding.
- To plan for schools, roads, and hospitals.
- To perform the population benchmarking used for weighting surveys and for other statistical reporting purposes.

Data Availability, Access, and Use

AVAILABILITY

The timeline for releasing data and the specific data products available vary from census to census. The 2020 Census results were released over a three-and half-year period, from April 2021 through September 2024. There are nine data products:

1. Apportionment.
2. Redistricting File (P.L. 94-171).
3. Demographic Profile.
4. Demographic and Housing Characteristics File (DHC).
5. Congressional District Summary Files.
6. Detailed Demographic and Housing Characteristics Files A (Detailed DHC-A).
7. Detailed Demographic and Housing Characteristics Files B (Detailed DHC-B).
8. Privacy-Protected Microdata File (PPMF).
9. Supplemental Demographic and Housing Characteristics File (S-DHC).

More information about each of these data products and how to access the data can be found on the Census Bureau's About 2020 Census Products webpage.

ACCESS

The Census Bureau releases decennial data in multiple formats, including as tables on data.census.gov, in the Census Bureau's application programming interface (API), and in other tools. The Census Bureau also provides descriptions of each of these data products.

The Census Bureau has produced a wealth of documentation about the decennial census methodology, its quality, and how to use the data.

USE

National KIDS COUNT does not actively use decennial census data for the *KIDS COUNT Data Book* or KIDS COUNT Data Center, as these rely on more regularly updated population data. However, the completeness and quality of census data are critical to all KIDS COUNT activities. For example, decennial population counts are a significant component of population weights for all federal surveys and population surveys.

On occasion, KIDS COUNT state programs may need to use the decennial counts. Example use cases include to understand the population size of smaller communities and in denominators for rates at substate geographic levels.

Questionnaire

The decennial census counts residents where they usually live and sleep. One person fills out the form and reports information for all household members.

The questionnaire includes a few questions about the household and each member of the household (see Figure 1). The 2020 Census asked about the number of people living in the household and whether the home is owned or rented. For each individual in the household, the form asked for the name, sex, age, race, Hispanic origin, and relationship to "Person 1"—the person completing the questionnaire and typically the owner or renter of the household.

The 2020 Census questions were mostly unchanged from the 2010 Census. The most notable changes were new write-in areas under the race categories for those who identify as white or Black to provide information on origin and new household relationship categories to identify same-sex and opposite-sex partners and spouses.

FIGURE 1. HOUSEHOLD QUESTIONS FROM THE 2020 CENSUS

Source: U.S. Census Bureau, "Decennial Census of Population and Housing Questionnaires and Instructions," revised November 23, 2021.

A Note on Disclosure Avoidance and Differential Privacy

The 2020 Census was the first to use differential privacy, also known as disclosure avoidance, as its method to protect respondent confidentiality. Differential privacy is a scientific framework for processing data to protect the identities and personal information of the people in the data. It works by adding statistical noise—small, random additions or subtractions—to published statistics so that no one can reidentify a specific person or household with any certainty using any combination of the published data.

PRB partnered with the Census Bureau to develop a new series of materials to provide concise, reader-friendly information to data users about the new disclosure avoidance methods in the 2020 Census. Data users should review the Census Bureau's resources on the disclosure avoidance system and differential privacy (see Figure 2). Keep in mind that the application of differential privacy and considerations for data users vary across the 2020 Census data products.

FIGURE 2. RESOURCES ON THE 2020 DECENNIAL CENSUS DISCLOSURE AVOIDANCE SYSTEM

Understanding Differential Privacy

Why we're modernizing protections – and how differential privacy works.

Disclosure Avoidance for the 2020 Census: An Introduction

In this handbook, the U.S. Census Bureau's Disclosure Avoidance System i…

2020 Disclosure Avoidance System: Frequently Asked Questions

Answers to questions about Disclosure Avoidance and the 2020 Census.

Disclosure Avoidance Protections by Data Product

Learn more about why we are modernizing protections and how…

2020 Census Briefs

Series of briefs describing how disclosure avoidance methods are bein…

Source: U.S. Census Bureau, "2020 Decennial Census: Processing the Count: Disclosure Avoidance Modernization," revised December 4, 2023.

U.S. Census Bureau Population Estimates Program

The Census Bureau's Population Estimates Program (PEP) produces **the official estimates of the population** for the United States, states, metropolitan and micropolitan statistical areas, counties, cities, and towns, as well as for Puerto Rico and its municipios. PEP also provides the official estimates of housing units for the nation, states, and counties.

Population estimates have many uses. They give us information about population size; selected characteristics of the population; and how demographic changes like births, deaths, and migration impact population size. These estimates are used in funding allocations, policy and program planning and implementation, survey controls, and denominators for vital rates.

Data Availability, Access, and Use

Each year, PEP publishes population estimates by age, sex, race, and Hispanic origin at the national, state, and county levels as a time series back to the last census. Note that age categories vary across geographic levels; for example, single year of age estimates are not available at the county level, and annual population estimates for cities or towns are only available for the total population. You can request a special tabulation of population data from the Census Bureau; however there may be an associated cost.

AVAILABILITY

Population estimates are released on a rolling basis, beginning with the national and state total estimates in December and ending with population estimates for the nation, states, and counties by demographic characteristics the following summer, usually in late June. The Census Bureau maintains an updated schedule (see Table 1, next page, for an example.)

ACCESS

Population estimates are released in several formats. The Census Bureau publishes formatted tables and datasets on the PEP webpage. Data are also available in CDC Wonder, though this release lags the release by the Census Bureau.

USE

Each annual release of the population estimates series is called a vintage, referring to the original production year. For example, Vintage 2023 was developed in 2023, with the first data release in December 2023. PEP uses a vintage for its annual estimates because it takes the most recent data on births, deaths, and migration and revises population estimates for all years since the latest decennial census (e.g., the 2020 Census). The current vintage of data supersedes all previously produced estimates for the included years. This means that Vintage 2023 includes estimates for 2020, 2021, 2022, and 2023, and these estimates should be used instead of prior-year vintages for earlier years of estimates.

On the KIDS COUNT Data Center, population estimates are updated annually using the most recent vintage year data. Each year, we replace all prior years of population estimates data back to the most recent decennial census. But for vital rates, such as birth and death rates, we match the vintage year with the birth or death year. We do not revise past-year vital rates with revised population estimates.

Postcensal and Intercensal Estimates

The population estimates that KIDS COUNT uses for calculating demographic indicators and vital rates are postcensal estimates. These are a time series of population estimates and population changes since the last census (e.g., Vintage 2023, as discussed on p. 10).

Following the decennial census, the Census Bureau adjusts the prior 10-year time series to smooth the transition from one decennial census to another with a set of **intercensal estimates**. Once produced, the intercensal estimates become the preferred series of data for the decade.

Therefore, if you need **a population estimate that falls between prior census years, you should use the intercensal estimates for those years**. For example, if you need a population estimate for a year between 2010 and 2020, you should use the 2010-2020 and intercensal estimates (scheduled for a rolling release beginning November 2024). If you need **a population estimate for a year since the last census, use the corresponding year in the most recent vintage of postcensal estimates**. For example, for a 2022 population estimate, use the 2022 data in the Vintage 2023 postcensal estimates.

TABLE 1. ANNUAL POPULATION AND HOUSING UNIT ESTIMATES

Population Detail	Latest Data Available	Date of Next Release
National and state population and demographic components of population change	July 1, 2023	December 2024 (July 1, 2024 data)
Puerto Rico population and demographic components of population change	July 1, 2023	December 2024 (July 1, 2024 data)
National, state, and Puerto Rico population age 18 and over	July 1, 2023	December 2024 (July 1, 2024 data)
County population and demographic components of population change	July 1, 2023	March 2025 (July 1, 2024 data)
Metropolitan and micropolitan statistical area population	July 1, 2023	March 2025 (July 1, 2024 data)
Puerto Rico municipio population	July 1, 2023	March 2025 (July 1, 2024 data)
National population by age and sex	July 1, 2023	April 2025 (July 1, 2024 data)
City and town (incorporated place and minor civil division) population	July 1, 2023	May 2025 (July 1, 2024 data)
National, state, and county housing units	July 1, 2023	May 2025 (July 1, 2024 data)
National population by age, sex, race, and Hispanic origin	July 1, 2023	June 2025 (July 1, 2024 data)
State population by age, sex, race, and Hispanic origin	July 1, 2023	June 2025 (July 1, 2024 data)
Puerto Rico population by age and sex	July 1, 2023	June 2025 (July 1, 2024 data)
Metropolitan and micropolitan statistical area population by age, sex, race, and Hispanic origin	July 1, 2023	June 2025 (July 1, 2024 data)
County population by age, sex, race, and Hispanic origin	July 1, 2023	June 2025 (July 1, 2024 data)
Puerto Rico municipio population by age and sex	July 1, 2023	June 2025 (July 1, 2024 data)

Source: U.S. Census Bureau, "Annual Population and Housing Unit Estimates Schedule," revised November 5, 2024.

American Community Survey

The Census Bureau's American Community Survey (ACS) is a nationwide survey that **collects information on the demographic, social, economic, and housing characteristics of the U.S. population**. Updated annually, it is the highest-quality source of local, comparable data on communities across the nation. Data analysts, government officials, businesses, and many others use these statistics for local planning and policy development and to track the population's well-being.

Each year a representative sample of 3.5 million addresses receives a survey, and information is collected nearly every day. Each annual release from the ACS contains data collected in the prior year, making it one of the most up-to-date resources on the U.S. population. Typically, data releases begin in September and continue through January of the next calendar year.

1-Year and 5-Year Estimates

The Census Bureau produces two main types of ACS estimates: 1-year estimates and 5-year estimates. The ACS 1-year estimates provide data for a 12-month calendar year period, and the ACS 5-year estimates combine 60 consecutive months of data to produce estimates that cover a five-year period.

The 1-year estimates provide data for geographic areas, such as counties and cities, with populations of 65,000 or larger. The 5-year estimates provide data for all geographic areas. Figure 3 shows U.S. counties

FIGURE 3. U.S. COUNTIES BY POPULATION SIZE, 2023

< 20,000   20,000–65,000   ≥ 65,000
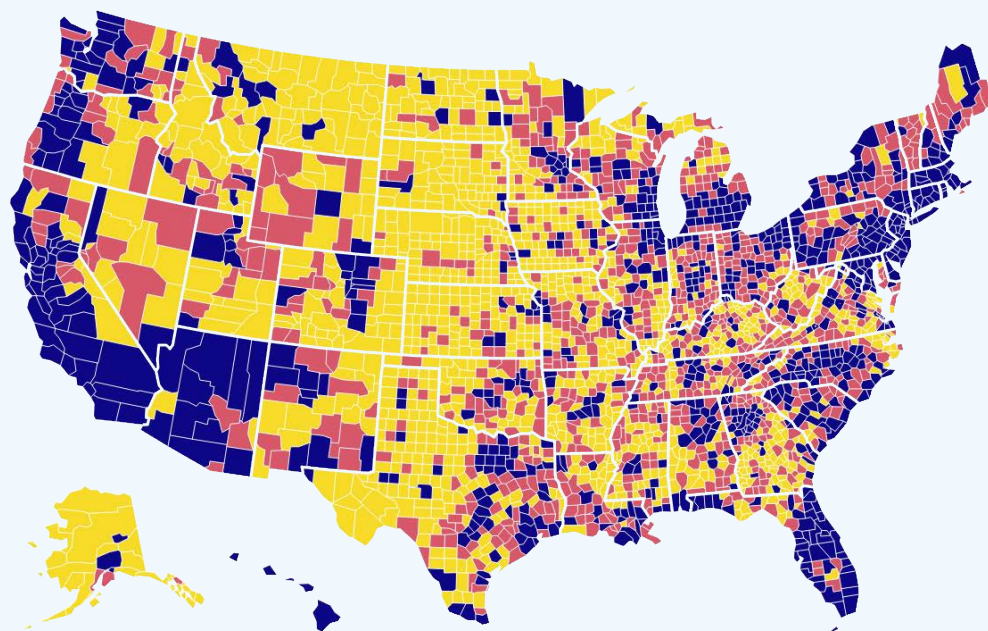
Source: PRB analysis of data from the U.S. Census Bureau, Population Estimates Program.

by size; those shaded dark blue are large counties with data available in the ACS 1-year estimates.

Whether to use 1-year or 5-year estimates depends on several factors, such as the size of the geographic area of interest, the importance of having the most current estimate, and the need for the most precise estimate. Table 2 (next page) provides some guidance on when you might use each set of estimates.

As a rule of thumb, if you need data for counties or subpopulation groups (by race and ethnicity, for example), you will likely need to use the 5-year estimates.

Geography

ACS data are tabulated for a variety of geographic areas, ranging from large areas such as the nation and states to smaller areas such as cities, school districts, and census tracts.

Many geographic areas are organized in a hierarchy, with smaller areas nested within larger areas (see Figure 4). For example, census tracts are nested within counties, and counties are nested within states. The wide range of available geographic areas allows data users to drill down to describe the characteristics of their communities, how they compare with other geographic areas, and how they have changed over time.

TABLE 2. DISTINGUISHING FEATURES OF ACS 1-YEAR AND 5-YEAR ESTIMATES

1-Year Estimates	5-Year Estimates
12 months of collected data	60 months of collected data
Data for areas with populations of 65,000+	Data for all areas
Smallest sample size	Largest sample size
Less reliable than 5-year estimates	Most reliable
Most current	Least current
Best used when: • Currency is more important than precision. • Analyzing large populations.	Best used when: • Precision is more important than currency. • Analyzing very small populations. • Examining tracts and other small geographic areas because 1-year estimates are not available.

Source: U.S. Census Bureau, "Using 1-Year or 5-Year American Community Survey Data," revised October 31, 2024.

FIGURE 4. HIERARCHY OF SELECT GEOGRAPHIC ENTITIES IN THE AMERICAN COMMUNITY SURVEY
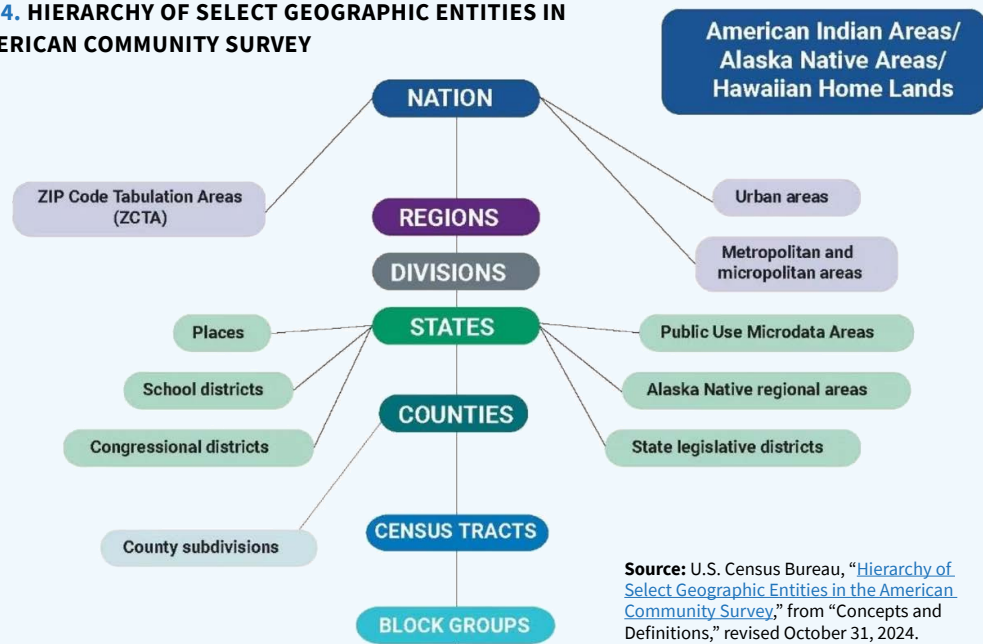
Source: U.S. Census Bureau, "Hierarchy of Select Geographic Entities in the American Community Survey," from "Concepts and Definitions," revised October 31, 2024.

Subjects

The ACS includes many topics that are relevant for child well-being, such as school enrollment and educational attainment, employment status, income, health insurance coverage, and family structure.

Subjects fall into four main categories: social, housing, economic, and demographic (see Table 3). Knowing the subject names and definitions is key to searching for data and understanding what is being measured.

TABLE 3. SUBJECTS IN THE ACS, BY CATEGORY

Social Subjects	Housing Subjects	Economic Subjects	Demographic Subjects
Ancestry	Bedrooms	Class of Worker	Age; Sex
Citizen Voting-Age Population	Computer and Internet Use	Commuting (Journey to Work) and Place of Work	Group Quarters Population
Citizenship Status	House Heating Fuel	Employment Status	Hispanic or Latino Origin
Disability Status	Kitchen Facilities	Food Stamps/SNAP	Race
Educational Attainment	Occupancy/Vacancy Status	Health Insurance Coverage	Relationship to Householder
Fertility	Occupants per Room	Income and Earnings	Total Population
Grandparents as Caregivers	Plumbing Facilities	Industry	
Language Spoken at Home	Rent	Occupation	
Marital History	Rooms	Poverty Status	
Marital Status	Selected Monthly Owner Costs	Work Status Last Year	
Migration/Residence 1 Year Ago	Telephone Service Available		
Place of Birth	Tenure (Owner/Renter)		
School Enrollment	Units in Structure		
Undergraduate Field of Degree	Value of Home		
Veteran Status; Period of Military Service	Vehicles Available		
Year of Entry into the U.S.	Year Householder Moved Into Unit		
	Year Structure Built		

Source: U.S. Census Bureau, "Subjects Included in the Survey," revised August 19, 2024.

ACS PUBLIC USE MICRODATA SAMPLE

The [American Community Survey Public Use Microdata Sample (ACS PUMS)](#) includes a subsample of the full ACS microdata that can be used to create custom tabulations and multivariate analyses that are not available in other existing products. ACS PUMS files contain individual records with information about each person and housing unit in the sample.

Because these files are for public use, they have been produced in a way that protects confidential information. These protections include:

- Using a subset (about two-thirds) of the full ACS sample.

- Excluding personally identifiable information, such as names and addresses.

- Using data processing techniques to swap or truncate extreme values.

- Limiting the geographic areas that can be identified in the sample.

Because of these changes, estimates derived from the ACS PUMS will not exactly match estimates in published tables on data.census.gov.

ACS PUMS data contain individual records that need to be weighted and aggregated for analysis, meaning they can be a bit more difficult to work with than published tables. Data users will need to use statistical programming software, such as STATA, SAS or R, or the [Microdata Access Tool](#) (MDAT) developed by the Census Bureau. The ACS PUMS handbook, [*Understanding and Using the American Community Survey Public Use Microdata Sample Files: What Data Users Need to Know*](#), provides detailed guidance on getting started with this source.
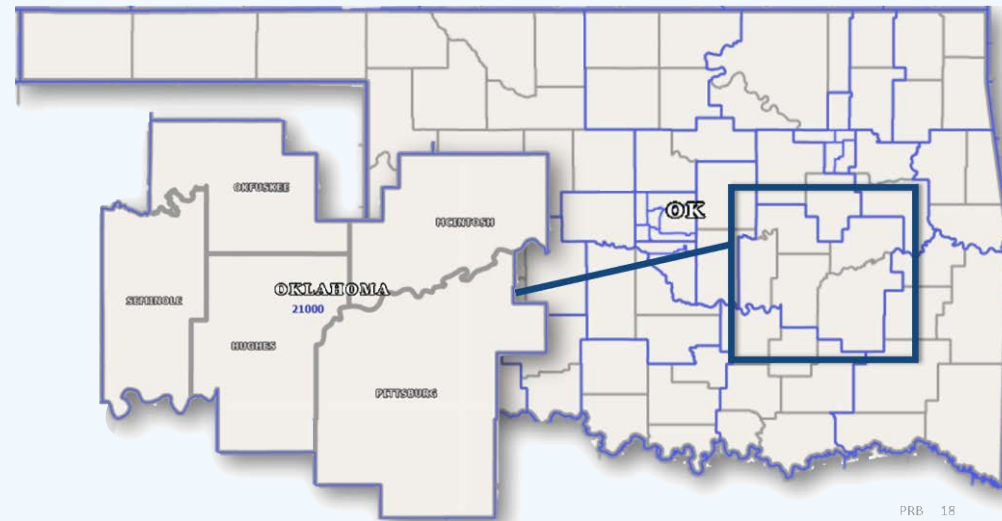
Source: PRB analysis of geographic boundaries using [TIGERweb](#).

Geographies in the ACS PUMS

Besides the nation as whole, geographies available through the ACS PUMS include states and Public Use Microdata Areas (PUMAs). PUMAs are statistical geographic areas that have a minimum population of 100,000. PUMAs use census tracts as their building blocks and do not cross state lines. You can use PUMAs to indirectly derive substate geographies, such as counties. Typically, counties with especially large populations are subdivided into two or more PUMAs, while PUMAs in more rural areas are made up of groups of adjacent, less-populous counties.

Note that PUMA boundaries are redrawn every 10 years following the decennial census. When using

PUMAs with ACS 5-year data and drawing on PUMA boundaries from two different vintage years, you need to make sure the boundaries are consistent across datasets. Another option is to pool larger geographic areas together until you have a consistent land area across the two vintages.

Two tools are especially useful for linking PUMAs to other substate geographies. [TIGERweb](#) is a mapping tool that visualizes geographic relationships. In TIGERweb, you can see how the PUMA boundaries (blue lines) and county boundaries (gray lines) intersect, among many other geographic relationships. Figure 5 shows county lines in Oklahoma (gray lines) and how counties are grouped together into corresponding

TABLE 4. GEOCORR RESULTS LINKING PUMAS TO FIVE OKLAHOMA COUNTIES

State Code	PUMA (2022)	County Code	State Abbreviation	County Name	PUMA22 Name	Total Population (2020 Census)	County-to-puma22 Allocation Factor	Puma22-to-county Allocation Factor
40	21000	40063	OK	Hughes OK	East Central Oklahoma Counties--McAlester, Seminole, Holdenville & Eufaula Cities	13,367	1	0.1205
40	21000	40091	OK	McIntosh OK	East Central Oklahoma Counties--McAlester, Seminole, Holdenville & Eufaula Cities	18,941	1	0.1707
40	21000	40107	OK	Okfuskee OK	East Central Oklahoma Counties--McAlester, Seminole, Holdenville & Eufaula Cities	11,310	1	0.1019
40	21000	40121	OK	Pittsburg OK	East Central Oklahoma Counties--McAlester, Seminole, Holdenville & Eufaula Cities	43,773	1	0.3945
40	21000	40133	OK	Seminole OK	East Central Oklahoma Counties--McAlester, Seminole, Holdenville & Eufaula Cities	23,556	1	0.2123

Source: PRB analysis of Geocorr 2022: Geographic Correspondence Engine, Missouri Census Data Center.

PUMAs (blue lines), with a detailed example showing the grouping of Okfuskee, McIntosh, Seminole, Hughes, and Pittsburg counties into a single PUMA, 21000.

Geocorr, a web-based application from the Missouri Census Data Center, allows users to create Excel-based geographic correspondence files that quantify the relationship between two or more geographic areas. This tool uses an allocation factor to show the portion of a source area (in terms of land mass or population), such as a PUMA, that corresponds with the target area, such as a county. An added benefit of Geocorr is that it provides the FIPS and PUMA codes you'll need to tabulate your data with a microdata tool.

Using the above example from Oklahoma (see Table 4), Geocorr results show that each of the five examined counties falls entirely within a single PUMA (county-to-puma22 allocation factor equals 1) and that no other counties contribute to that PUMA (summing the puma22-to-county allocation factor across the counties equals 1).

Data Availability, Access, and Use

AVAILABILITY

ACS data are released each year, in the year following data collection. Data releases occur on a rolling basis, starting in mid-September with the ACS 1-year data, following in mid-to-late October with the 1-year PUMS, and ending in mid-December with the ACS 5-year data. The full data release schedule is available online.

The release schedule allows the data to be very timely—for example, data collected between January 1 and December 31, 2023, became available less than a year later, in September 2024.

ACCESS

ACS data are released across several data platforms or tools. Based on your needs, you may be best served by using the Census Bureau's general data website (data.census.gov), the MDAT tool, the API tool, or the ACS Summary Files (see Figure 6, next page).

Data.Census.Gov

Data.census.gov is the main system for accessing and disseminating Census Bureau data online. The website offers pre-tabulated data and maps from the ACS and many other data sources, along with several types of data products and mapping tools (see Table 5).

In general, **profiles** offer a broad look at a community's social, economic, housing, and demographic characteristics. Profile tables focus on geography or population group, with data presented across many variables.

TABLE 5. ACS DATA PRODUCTS WITH YEARS AVAILABLE AND INCLUDED GEOGRAPHIES

Topics/Product Type Included	Year(s) Available & Data Set(s) Included	Geographies Included
Comparison Profiles, Data Profiles, Detailed Tables, Subject Tables	2006-2010 ACS 5-year estimates to latest release	All areas down to block group level
Data Profiles, Detailed Tables	2006-2010 and 2011-2015 ACS 5-year Selected Population Tables and American Indian and Alaska Native Tables	Selected geographies down to tract level
Supplemental Tables	2014 ACS 1-year Supplemental Estimates to latest standard release	Areas with populations of 20,000+
Comparison Profiles, Data Profiles, Detailed Tables, Selected Population Profiles, Subject Tables	2010 ACS 1-year estimates to latest standard release	Areas with populations of 65,000+

Source: U.S. Census Bureau, "American Community Survey Data Tools," revised October 9, 2024.

The **Selected Population Profiles** provide social, economic, and housing data disaggregated by race, ethnicity, ancestry, and county/region of birth. They are produced annually with the 1-year estimates and provide data for larger population areas.

Subject tables and detailed tables provide a more comprehensive view of a subject. Many tables are available with disaggregation by age, sex, race, ethnicity, and other characteristics. These tables may be a little more challenging to work with because of the frequent need to combine estimates to calculate percentages or rates.

For example, to find the number of children under age 12 living below the federal poverty line, you would need to use table B17001, and add males below the poverty lines at ages "under 5 years," "5 years," and "6 to 11 years" with females in the comparable categories (i.e., the numerator). To calculate the percentage of children under age 12 living below the federal poverty line, you would need to add the comparable categories for males and females living above the federal poverty line to the previously calculated numerator to get the denominator, divide the numerator by the denominator, then multiply by 100.

The Census Bureau [provides many resources](#) on how to use data.census.gov. More advanced users also have several options to access detailed ACS data through downloadable Summary Files, the MDAT tool, ACS PUMS files, or the Census Bureau's API (see Figure 6).

MDAT

[MDAT](#) is a web-based tool that can create custom tables with ACS PUMS data. Using the point-and-click interface, you can create a data table with the parameters of your choice. There are [many webinars, videos, and guides](#) on how to use this tool. Using MDAT requires understanding how the ACS PUMS data are different than the pre-tabulated data. More information on the ACS PUMS is available in this guide and on the Census Bureau's website.

API

The Census Bureau's [API](#) is a tool that allows users to access ACS data efficiently. You can use it to create simple queries to access ACS data through an internet browser or write a set of commands to access ACS data via your own programs or applications. With the API, you can also develop standardized data pulls that are easily updated year after year, eliminating the need to manually download new data from data.census.gov. Learn more with the *Census Data API User Guide*.

Summary Files

The ACS [Summary Files](#) include all estimates and margins of error (MOEs) from the Detailed Tables for all geographies. With these data, you

FIGURE 6. ACS DATA ACCESS TOOLS, BY SKILL LEVEL OF DATA USER

Data. census.gov — MDAT — Census Data API — ACS Summary File

NOVICE   ADVANCED
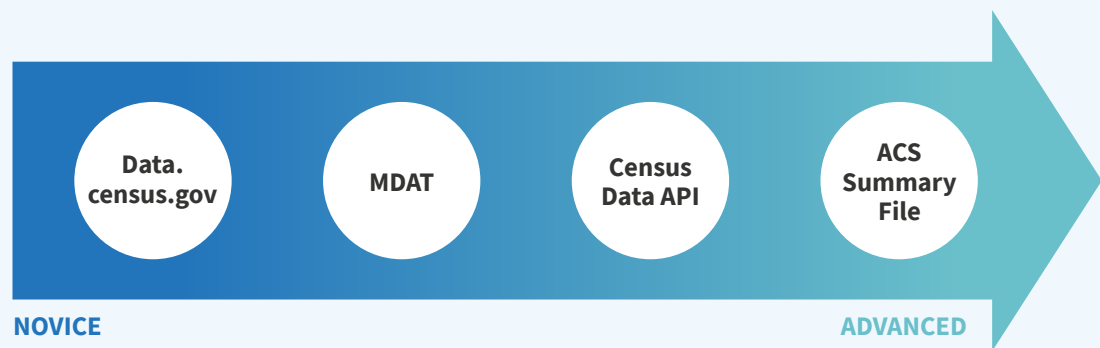
Source: PRB and U.S. Census Bureau visualization of ACS data access tools.

can download statistics for many tables and geographic areas. Using these data files requires advanced knowledge of the ACS and the help of a statistical software program, such as SAS.

USE

ACS data have many uses. Data analysts, government officials, businesses, and many others use these statistics for local planning and policy development and to track the population's well-being. The ACS was designed to provide estimates of the **characteristics** of the population. This is different from the Population Estimates Program, which is designed to provide **counts** of the population across key subgroups.

ACCURACY

Because the ACS is based on a sample, ACS data are estimates. All estimates from sample surveys have a degree of uncertainty associated with them called sampling error. Sampling error is the estimated difference between an estimate based on a sample and the actual (but unknown) values based on the entire population. It is used to measure the precision or reliability of estimates.

When more of a population is included in a sample, it's more likely to be representative of the entire population. Therefore, in general, the larger the sample—larger geographic areas, for example—the smaller the sampling error.

Table-Based Estimates

The ACS tables released on data.census.gov and other platforms provide a 90% margin of error (MOE) for published numbers and percents. The 90% MOE tells you to what degree the derived estimates may differ from the true population values. Specifically, it means that 9 out of 10 times, the true population value will fall between the sample estimate and plus or minus the MOE.

The Census Bureau provides formulas for calculating the MOE when combining categories or calculating percentages. One place you can find these formulas is in the ACS handbook, _Understanding and Using American Community Survey Data: What All Data Users Need to Know_.

In the KIDS COUNT Data Center, PRB uses the MOEs for percents to determine which estimates are reliable and usable. We suppress estimates that have an MOE larger than 5 percentage points, meaning the confidence interval (CI) is 10 percentage points or wider.

FIGURE 7. ACS DATA USERS GROUP WEBSITE
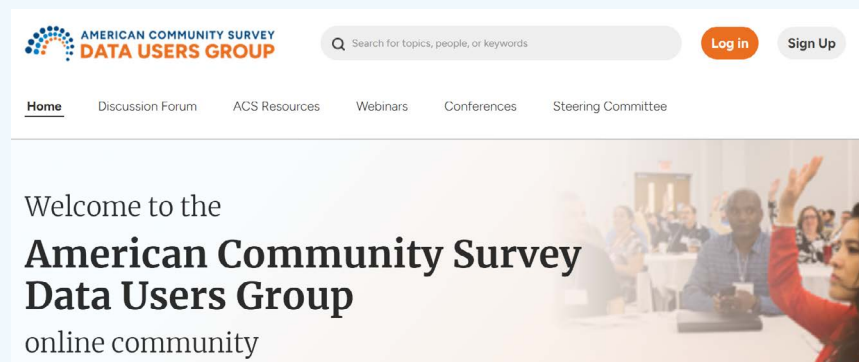
Source: American Community Survey Data Users Group, "Welcome to the ACS Data Users Group Online Community."

For more information and examples on how to calculate MOEs, CIs, and standard errors (SEs) from estimates available through data.census.gov (including percentages), please see the annually updated Accuracy of the Data document.

More information on assessing the reliability of estimates across data sources is below.

Estimates From ACS PUMS

Estimates from the ACS PUMS will have larger errors than the same estimates drawn from the full ACS. This is because the estimates are from a smaller sample.

Data users need to calculate standard errors (SEs) for estimates derived from the ACS PUMS. The ACS PUMS handbook describes how to calculate the SE using two methods: the generalized SE formula and the successive difference replication formula. The successive difference replication formula will provide much more accurate SEs but is more challenging to implement.

FURTHER LEARNING

The Census Bureau provides many resources to assist data users with using and understanding the ACS. The resources are available in many formats, including videos, webinars, handbooks, and technical documents.

The ACS Data Users Group promotes information sharing among data users about key ACS data issues through an online community, webinars, and a biennial conference. In the online community, users can share messages, materials, and announcements related to the ACS (see Figure 7). Membership is free and open to all interested ACS data users. The ACS Data Users Group is powered by PRB, in collaboration with the U.S. Census Bureau, with PRB staff and Census Bureau officials as active contributors to the online community.

Current Population Survey

The Current Population Survey (CPS) is a monthly survey sponsored by the Census Bureau in collaboration with the Bureau of Labor Statistics. It is the primary source for national labor force statistics, including the unemployment rate. It also provides a comprehensive body of data on other important economic characteristics, including poverty, earnings, work, and education.

Data Topics and Supplements

Monthly CPS

The monthly or basic CPS covers information about employment, unemployment, hours of work, earnings, and people not in the labor force. It is the official source for national estimates of the civilian labor force and unemployment rates.

Annual Social and Economic Supplement

The Annual Social and Economic Supplement, also known as the March Supplement, is the official source for national income and poverty data. Collected each March, it contains annual data on a variety of other social and economic measures, such as work experience, health insurance coverage, noncash benefits, and program participation.

Supplements

Each CPS also includes a supplemental questionnaire on a specific topic. Topics vary month to month, with certain topics surveyed annually, others every other year, and others only once. Topics include child support, volunteerism, food security, school enrollment, computer and internet use, and voting, among others.

Data Availability, Access, and Use

AVAILABILITY

CPS data are released throughout the year. Basic monthly CPS data are released the month after data collection. Annual Social and Economic Supplement data are released in September following data collection in the previous March. CPS supplements are released anywhere from six to 18 months after data collection.

ACCESS

Monthly CPS

Data are available through multiple tools on the Bureau of Labor Statistics website. You can use the complete microdata files or create custom tabulations using MDAT.

Annual Social and Economic Supplement

New data for this supplement are released each September. Results are reported in the Census Bureau's annual reports on income and poverty. You can access the data through the pretabulated data tables and microdata files, or create custom tabulations with MDAT.

Supplements

The Census Bureau provides information on all CPS supplements and what years of data are available (through 2021), as well as a more detailed list about each topic and the available microdata. You can also use MDAT for custom tabulations of CPS supplement data.

USE

CPS data provide reliable estimates for the nation, but the Census Bureau warns that one-year CPS estimates may be unreliable for many states. To produce reliable state-level estimates, data users generally need to create two- or three-year averages. The sample size does not allow reliable estimates to be obtained at the county level. In fact, not all counties are included in the sample, and data are not available for most counties that are sampled due to confidentiality laws.

Tabulating by race and ethnicity at the state level requires pooling at least three years of data and will likely only produce reliable estimates for large states.
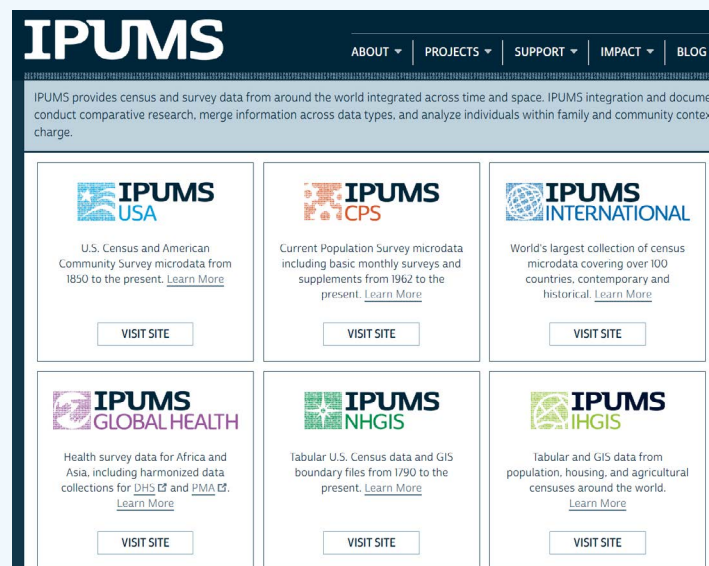
IPUMS

IPUMS used to be an acronym for the Integrated Public Use Microdata Series. Today, it's just IPUMS and is a center at the University of Minnesota working to make social and economic data more accessible. It does this by integrating data across time and space, publishing user-friendly documentation, and developing data dissemination tools.

IPUMS is a robust resource for learning about data. The website includes documentation about survey samples, information on variables (including what's being measured and how it's changed over time), various methodological details, and guidance for using IPUMS tools (see Figure 8).

IPUMS provides census and survey data from around the world, including a vast amount of data from the U.S. Census Bureau. For KIDS COUNT, key data resources are IPUMS USA, IPUMS CPS, and IPUMS NHGIS. IPUMS USA provides decennial census and ACS data; IPUMS CPS provides data from the basic monthly CPS and all CPS supplements; and IPUMS NHGIS provides tabulated decennial census and ACS data, with geographic shape or correspondence files attached.

A benefit of these resources is that IPUMS has worked to harmonize data within data sources, incorporating survey or geography changes so that long-term time trend data are available. A caution with using these resources is that there are places where IPUMS implements its own algorithms to make changes to the original data source.

FIGURE 8. IPUMS HOMEPAGE

Source: IPUMS Center for Data Integration, "IPUMS", Institute for Social Research and Data Innovation, University of Minnesota.

When using IPUMS, be aware that:

- IPUMS variables may differ from the corresponding Census Bureau versions. Family structure and family relationships, family income, poverty, and race are some examples.
- Estimates using the IPUMS versions will not always match estimates from the original data source.
- Variable levels, codes, and labels may not match those from the original data source.

To use IPUMS, first select the data source (e.g., USA, CPS, NHGIS, etc.). Then, within the data source, select the tools or documentation you need. You can learn about the variables in the data source (both from the perspective of the original source and any changes by IPUMS), download microdata files, and even create custom tabulations with the online data analysis tool.

National Vital Statistics System

The National Vital Statistics System is a coordinated effort by the U.S. Centers for Disease Control and Prevention (CDC), the National Center for Health Statistics, and the states to provide the most complete data on births and deaths in the United States. Vital registration systems operated in all states and territories submit data to this system.

National KIDS COUNT compiles data on births overall and by selected maternal and infant risk factors. Mortality indicators include infant mortality, child and teen deaths, and teen deaths by selected causes.

Vital statistics data are publicly available for the nation, states, and the District of Columbia. Birth data are available for large counties with at least 100,000 people, and death data are available for all counties. Data for U.S. territories are also available through published tables and microdata.

Data Availability, Access, and Use

AVAILABILITY

Data on vital events, such as births and deaths, are released monthly as a provisional count. Final birth and death data are released annually. Birth data for the prior year are typically available in the fall, and death data are typically released in December.

ACCESS

Vital statistics data can be accessed in multiple ways. CDC Wonder can be used to create custom tabulations, and published tables are available through National Center for Health Statistics publications. You can also use public use microdata files, though because these do not contain geographic information, they can only support analysis at the national level or for each U.S. territory.

The CDC also maintain WISQARS, a set of online analysis tools for data on fatal and nonfatal injuries and related costs.

State KIDS COUNT organizations may prefer to use data from their own state departments of health or vital records offices.

USE

Vital statistics are a core component to developing population estimates and are used for planning purposes. They are also critical to understanding public health.

In addition to providing counts of births and deaths, vital statistics data include information across key characteristics. Data users can analyze birth records by age, race, ethnicity, nativity, and education. You can also examine maternal characteristics, including pregnancy history, prenatal care characteristics, and maternal and pregnancy risk factors; labor, delivery, and infant characteristics; and selected paternal characteristics. Death records provide information on age, gender, race, ethnicity, place of death, and leadings cause(s) of death.

National Survey of Children's Health

The National Survey of Children's Health (NSCH) provides information on the physical and emotional health of children ages 0 to 17. In addition to assessing health outcomes, the survey also examines other factors related to children's well-being, such as access to and quality of health care, family relationships, neighborhood characteristics, and school experiences. The NSCH is also designed to measure the prevalence and impact of special health care needs among U.S. children.

The survey is sponsored by the Maternal and Child Health Bureau of the U.S. Health Resources and Services Administration. More information about the topics, data, and history of the NSCH can be found at the Child and Adolescent Health Measurement Initiative's Data Resource Center for Child and Adolescent Health.

Data Availability, Access, and Use

AVAILABILITY

NSCH data are released annually, usually in the fall the year after data collection. For example, data collected in 2023 were released in fall 2024.

ACCESS

Data are released in public use microdata files and as tables in an interactive query system.

USE

Data are available at the national and state levels and can be further disaggregated by race and ethnicity, income, and other key demographic and health characteristics.

One-year data can be used for national estimates, two-year data can provide estimates for states in total, and five-year data can provide estimates for states among subpopulations. Some states have sponsored oversamples; for these states, it may be possible to use fewer years of data. Please review technical documentation about any oversamples for guidance on using the data.

PRB can provide technical assistance with NSCH data. However, Child Trends, a nonprofit research center, is the main NSCH data provider for the national KIDS COUNT Data Center.

Small Area Income and Poverty Estimates

VIEW SOURCE →

The Census Bureau's Small Area Income and Poverty Estimates (SAIPE) program produces single-year estimates of income and poverty for all U.S. states and counties as well as estimates of school-age children living in poverty for the country's 13,000+ school districts. The program is used to allocate billions of dollars in federal funding annually to school districts and small localities. State and local programs use these estimates to distribute funds and manage programs.

SAIPE estimates are produced by combining data from several sources, including the ACS, population estimates, and administrative data sources.

Data Availability, Access, and Use

AVAILABILITY

SAIPE data are updated annually, with new data available a few months after the ACS 1-year estimates are released. For example, ACS 1-year estimates are released in September, and the corresponding SAIPE data are released in December.

ACCESS

SAIPE data can be accessed through multiple tools, including an API, datasets, interactive data tools, and data tables.

USE

State and county estimates are available for:

- Total number of people in poverty.
- Number of children under age 5 in poverty (for states only).
- Number of related children ages 5 to 17 in families in poverty.
- Number of children under age 18 in poverty.
- Median household income.

School district estimates include:

- Total population.
- Number of children ages 5 to 17.
- Number of related children ages 5 to 17 in families in poverty.

Small Area Health Insurance Estimates

The Census Bureau's Small Area Health Insurance Estimates (SAHIE) program is the only source of data for single-year estimates of health insurance coverage status for all U.S. counties by selected economic and demographic characteristics.

The program produces its estimates by combining data from several sources, including the ACS, population estimates, and administrative data sources.

Data Availability, Access, and Use

AVAILABILITY

Data are updated annually. The 2023 SAHIE estimates will be published in summer 2025.

ACCESS

SAHIE data are available through multiple tools, including an API, datasets, and an interactive data tool.

USE

SAHIE data are recommended for examining single-year estimates of health insurance coverage for all U.S. counties. These estimates are more current than the ACS 5-year estimates, and time trend data are available beginning in 2000.

The SAHIE program publishes state and county estimates of the insured and uninsured population for the full cross-classification by selected age groups, sex, and income-to-poverty ratio, and by race and ethnicity (at the state level only). Estimates of insured and uninsured children ages 0 to 18 overall and by income-to-poverty ratios also are available for states and counties.

National Center for Education Statistics

The National Center for Education Statistics is a federal statistical agency of the U.S. Department of Education. The center collects, analyzes, and reports data on the state of American education. Many surveys and tools can help users gain a better understanding of this data.

Digest of Education Statistics

The Digest of Education Statistics is a data dissemination tool that contains a set of tables covering many aspects of American education from preschool through graduate school. The data tables are updated annually as data from each source become available.

National KIDS COUNT uses these tables to provide statistics on high school students not graduating on time and per pupil educational expenditures. These estimates may vary from those obtained by state departments of education because some states use definitions that vary slightly from the federal reporting standards.

Common Core of Data

Common Core of Data is a database on public elementary and secondary education. This database can be used to find information about schools at the state, district, and individual school levels. National KIDS COUNT uses these data for information on the number of schools and school enrollment, though additional data, such as revenue and expenditures, are available. Data can be accessed through an online table generator.

National Assessment of Educational Progress

National KIDS COUNT uses data from the National Assessment of Educational Progress to provide comparable state-level data on student achievement. While the program encompasses many grades and subjects, KIDS COUNT focuses on fourth grade reading and eighth grade math data. Assessments occur every other year. Schools are typically surveyed in the winter, and results are released the following fall.

Data are released in the Nation's Report Card series, via an interactive data tool (Data Explorer), through an API, and in a restricted-access microdata file.

Adoption and Foster Care Analysis and Reporting System

The KIDS COUNT Data Center includes several indicators on children in foster care from the Adoption and Foster Care Analysis and Reporting System. State agencies are required to report to this system case-level information on all children in foster care and children who have been adopted.

Child Trends provides these data to the KIDS COUNT Data Center.

VIEW SOURCE →

National Child Abuse and Neglect Data System

The KIDS COUNT Data Center includes indicators from the National Child Abuse and Neglect Data System. The system gathers information from all 50 states, the District of Columbia, and Puerto Rico about reports of child abuse and neglect. States voluntarily submit their data annually.

Child Trends provides these data to the KIDS COUNT Data Center.

VIEW SOURCE →

Census of Juveniles in Residential Placement

The Census of Juveniles in Residential Placement provides information on youth in residential placement, such as detention centers. Every other year, these facilities are asked to provide data on the gender, age, race, and details about the offense for each young person on the census reference date, or the specific date on which the information was collected.

Data are collected every other year (odd numbers) and released two years later through the Easy Access to the Census of Juveniles in Residential Placement platform.

VIEW SOURCE →

National Survey on Drug Use and Health

The National Survey on Drug Use and Health provides national-, state-, and substate-level estimates of substance use and mental illness. The survey is conducted each year, with updated estimates available annually. Data are released in a variety of formats, including reports, detailed tables, a custom tabulation tool, and microdata files.

State-level estimates are available for a limited number of indicators and require pooling two years of data together. These estimates are released annually through data tables and an interactive data tool. The data tables are separated into a set of prevalence estimates and a set of estimated counts, both with 95% CIs.

VIEW SOURCE →

KIDS COUNT
Data Book

16 Key Indicators of Child Well-Being

The KIDS COUNT index is comprised of 16 key indicators of child well-being split across four domains, which are then aggregated into a summary index. The index is updated and released each year as part of the *KIDS COUNT Data Book*.

Table 6 lists the data sources for each of these indicators.

TABLE 6. **INDICATORS IN THE KIDS COUNT DATA BOOK, BY DOMAIN**

INDICATOR	SOURCE
Economic Well-Being	
Children in poverty	American Community Survey, 1-yr tables (data.census.gov)
Children whose parents lack secure employment	American Community Survey, 1-yr microdata
Children living in households with a high housing cost burden	American Community Survey, 1-yr microdata
Teens not in school and not working	American Community Survey, 1-yr tables (data.census.gov)
Education	
Young children (ages 3 and 4) not in school	American Community Survey, 5-yr tables (data.census.gov)
Fourth graders not proficient in reading	National Assessment of Educational Progress
Eighth graders not proficient in math	National Assessment of Educational Progress
High school students not graduating on time	EdFacts file from the Digest of Education Statistics
Health	
Low birth-weight babies	National Vital Statistics System, Birth Data
Children without health insurance	American Community Survey, 1-yr tables (data.census.gov)
Child and teen deaths per 100,000	National Vital Statistics System, Mortality Data
Children and teens ages 10 to 17 who are overweight or obese	National Survey of Children's Health
Family and Community Context	
Children in single-parent families	American Community Survey, 1-yr tables (data.census.gov)
Children in families where the household head lacks a high school diploma	American Community Survey, 1-yr microdata
Children living in high-poverty areas	American Community Survey, 5-yr Summary Files
Teen births per 1,000	National Vital Statistics System, Birth Data

Source: PRB.

Best Practices for Data Use

Calculating Reliability

Much of the data for KIDS COUNT comes from surveys and only reflects a sample of the total population. All estimates from surveys have uncertainty associated with them—that is, the estimates derived from a sample survey will differ from the values that would have been obtained had the entire population been surveyed.

Measures of reliability, such as standard error (SE), provide an indication of the level of sampling variability in a dataset. SEs can then be used to develop other measures of reliability that are easier to interpret, such as margins of error (MOEs) and confidence intervals (CIs). These measures provide information on the likelihood of the survey estimate being within a certain range of the unknown population value. In general, smaller SEs mean more reliable or precise estimates.

The published ACS tables provide 90% MOEs for nearly all estimates. These MOEs can be used on their own, to calculate 90% CIs, or to calculate the SE from the MOE. Many other data sources will also provide an SE, an MOE, or a CI with published estimates. The steps below are applicable when measures of reliability are provided by the data source.

Calculating the CI Using the MOE

A CI is expressed as the range of values between a lower confidence bound and an upper confidence bound. Calculate an estimate's confidence bounds as follows:

Lower confidence bound = estimate – MOE

Upper confidence bound = estimate + MOE

The final CI can be expressed in different ways depending on your needs. A common example is as a range alongside the point estimate—e.g., 17% (15.1%, 18.9%).

Calculating the SE Using the MOE

Backing out the SE can be useful if you want to use a different level of significance (e.g., 95%), or if you want to calculate the coefficient of variation (p. 33). To calculate the SE from an MOE, you would divide the MOE by the z-score associated with the current confidence level. (The z-score is the critical value that defines the boundaries of the confidence interval.)

For example, to find the SE from a 90% MOE:

$$SE = \frac{MOE}{1.645}$$

Calculating the MOE When Combining Estimates From Published Tables

Many of the indicators in KIDS COUNT require combining categories across subgroups or calculating percentages from number data. This requires using formulas provided by the Census Bureau.

The formulas and how to use them are discussed in the ACS handbook as well as in the Accuracy of the Data statements that accompany each ACS data release. Please note that these formulas result in approximate SEs and MOEs because they do not consider correlations between estimates.

The discussion and formulas below focus on the ACS, but they can be used for any data source where the SEs or MOEs are provided.

Note that MOE and SE can be used interchangeably in the following formulas.

Calculating the MOE for Numbers When Combining Estimates Across Groups or Geographies

1. Get the margin of error for each of the number estimates: X_1, X_2,…

2. Square each MOE.

3. Sum the squared MOEs.

4. Take the square root of the sum.

$$MOE(X_1 + X_2) = \sqrt{(MOE_{X1})^2 + (MOE_{X2})^2}$$

Calculating the MOE for Percentages

Percent or ratio estimates require a different formula to calculate the MOE.

1. Follow the appropriate steps to obtain the MOE for the numerator (N) and denominator (D). (For example, use the published MOEs if applicable, or calculate the MOE by combining estimates as described above.)

2. Use the formula below to calculate the MOE for a percentage, p.

$$MOE(p) = 100 * \frac{1}{D} \sqrt{(MOE_N)^2 - \left[\left(\frac{N}{D}\right)^2 * (MOE_D)^2\right]}$$

Calculating the Coefficient of Variation

The coefficient of variation (CV) is a measure of reliability that shows the amount of variation in a sample relative to the mean. The CV measures how consistently close the values in the sample are to the mean of the dataset, with a smaller CV indicating higher consistency in the data.

$$CV = \frac{SE}{estimate} \left[\begin{array}{l} *100 \text{ to express} \\ \text{as a percentage} \end{array}\right]$$

The higher the CV, the more data observations are widely spread out around the mean. In other words, there is a lot of variation in the sample. A low CV, on the other hand, means that the data points are clustered closely around the mean. In other words, the dataset has low variability.

There are two disadvantages to using the CV. First, the CV is dependent on the magnitude of the estimate. For percent data, when the estimated percentage is close to 0 or close to 100, the CV will not provide meaningful results. Second, the CV is sensitive to outliers. If there are many extreme values in the dataset, the CV may not be useful.

Assessing Reliability

Once you've calculated measures of reliability for your estimates, you'll need to decide which information you want to share. In other words, when should estimates be suppressed because they are unreliable?

There are **no hard and fast rules**. Which data to show depends on the application.

National KIDS COUNT follows a few specific rules. First, always refer to the data source and follow any guidelines provided by the data provider. Then use other rules of thumb, such as considering data to be reliable when:

- CIs for percents are less than 10 percentage points wide (MOE<5).
- CVs are smaller than 30%.

Working With Small Areas or Small Populations

Small areas and small populations face several challenges with providing data. Frequently, the population may be so small that releasing data could potentially disclose confidential personal information. Other times, the measures of reliability are large; this happens because it is difficult to survey a small population without a huge sample size.

Borderline Reliable Estimates

Borderline reliable estimates are those that fall outside the previously described boundaries (e.g., CI widths that are larger than 10 percentage points or CVs larger than 30%), but are relatively close to these levels (e.g., CI widths that are 12 percentage points wide). To make these estimates more usable, you can:

- Present the estimate with a note to use with caution or add something like a red light/green light symbol.
- Use data aggregation techniques to increase the sample size and thus increase reliability.

Unreliable Estimates

Unreliable estimates have very large CVs or MOEs. If you still want to use the data, you can:

- Aggregate to increase sample size, if applicable.
- Suppress the estimates and highlight the need for more or better data.

AGGREGATION TECHNIQUES

There are multiple techniques to increase sample size. The choice of method depends on the data that are available to you as well as your goals.

Expanding the Geographic Area

Combine smaller levels of geography into a larger group (e.g., combine counties to create county groups).

Combining Multiple Years of Data

Pool multiple years of data together. Or, use ACS 5-year estimates instead of ACS 1-year estimates.

Combining Groups or Collapsing Categories

You may also consider combining subpopulations to create larger groups (e.g., combine ages 5 to 9 and ages 10 to 14 into ages 5 to 14). This may not always be a good option because it will mask variation across subpopulations, making it appear as though different populations have similar experiences. You will need to assess whether combining groups of people is appropriate.

Comparing Estimates and Measuring Change Over Time

Making comparisons requires ensuring that groups and geographies are comparable (see Figure 9). Specific to the ACS, the Census Bureau recommends using estimates to compare and describe population characteristics using the same time period and period length and testing for statistical significance. The Census Bureau also provides a tool that lists which ACS estimates can be compared.

Trend analysis requires taking into consideration additional factors that might make certain datasets unsuitable for comparison. Some things to consider include:

- **Geographic boundaries:** Such as changes in the boundaries of counties, cities, census tracts, or other geographic areas. These types of changes occur most often following a decennial census, but boundaries can change at other times as well.

- **Question and data processing changes:** For example, recent changes in the collection and processing of race and ethnicity data in the 2020 Census and 2020 ACS. Questionnaires are regularly reviewed and revised to incorporate changes in best practices in survey methodology and to respond to the changing needs of the population.

- **Major demographic changes:** Such as unequal increases or decreases in births, deaths, and migration across subgroups. Such shifts can alter the composition of who may be captured within population groups.

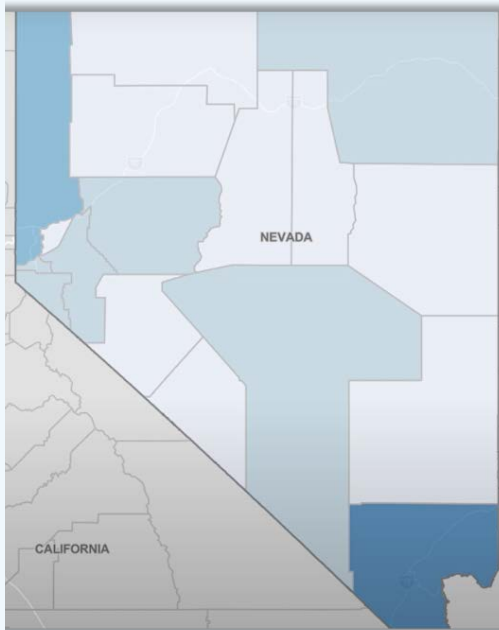**FIGURE 9.** DOS AND DONTS FOR MAKING COMPARISONS WITH ACS DATA

Use ACS for population characteristics (i.e., percent, means, medians, rates)

Compare non-overlapping datasets (i.e., 2012-2016 to 2017-2021)

Compare similar period lengths (i.e., 1-year to 1-year)

Conduct statistical testing when making comparisons between estimates

Use ACS for population totals (instead use decennial census or the Population Estimates Program)

Compare overlapping datasets (i.e., 2016-2020 to 2017-2021)

Compare estimates from different period lengths (i.e., 1-year to 5-year)

Look at estimates to decide if they are higher or lower than one another

Source: U.S. Census Bureau, *Understanding and Using American Community Survey Data: What All Data Users Need to Know*, 2020.

Statistical Significance

The Census Bureau has developed a simple way to conduct statistical testing without the use of statistical programming software. Their Statistical Testing Tool allows you to make comparisons between two groups or time periods.

You can also compare many groups, such as all the counties in your state (see Figure 10).

The Statistical Testing Tool is available online (see Figure 11). To use it, you need to have your estimates (numbers, percents, rates, etc.) and the corresponding margins of error. PRB worked with the Census Bureau to create a how-to video and a written guide, which provide detailed practical information on the tool and its features.

FIGURE 10. EXAMPLE FIGURE SHOWING VARIATION IN DATA ACROSS COUNTIES

Source: PRB visualization of an example of geographic variation.

FIGURE 11. SNAPSHOT OF THE STATISTICAL TESTING TOOL DEVELOPED BY THE U.S. CENSUS BUREAU

Source: U.S. Census Bureau, Statistical Testing Tool.

Disaggregating Data by Race and Ethnicity

Disaggregating data by race and ethnicity is important because it provides a **better understanding of strengths, needs, and quality of life** across different groups of people. A wide body of evidence has documented longstanding racial and ethnic disparities across many areas that are critical to health and well-being, including education, income, employment, health, and housing.

By disaggregating data, you can **identify, measure, and track racial and ethnic equity gaps** across many indicators. This process is important for uncovering and understanding the factors that contribute to racial and ethnic inequities.

Information on race and ethnicity is required for many federal programs and is key to **developing public policies that can better meet the needs of the population**. Accurate race and ethnicity information can help governments, organizations, and communities develop **programs that are better able to reach the people they're designed to serve** and that respect group differences.

MAJOR RACE AND ETHNICITY CATEGORIES

The U.S. Office of Management and Budget sets minimum categories that federal agencies must use when they collect information on race and ethnicity, a set of federal data standards known as Statistical Policy Directive No. 15 (SPD 15).

Implemented in 1997, the SPD 15 standards provide the required race and ethnicity categories that are currently available in data sources from the federal statistical system. These categories are broken into a two-question format, with one question on self-reported Hispanic or Latino origin (ethnicity) and another question on self-reported race. The race and ethnicity categories required by SPD 15 are as follows:

Ethnicity

- Hispanic or Latino
- Not Hispanic or Latino

Race

- American Indian or Alaska Native
- Asian
- Black or African American
- Native Hawaiian or Other Pacific Islander
- White

Respondents can select multiple races. A "two or more races" category can be used to report these responses.

Some data sources combine race and ethnicity categories into a single, mutually exclusive characteristic for reporting purposes; one example combination is Non-Hispanic American Indian or Alaska Native. Other sources report race and ethnicity separately.

More detailed race information may also be available for your reporting purposes. The ACS and the decennial census both provide checkboxes and write-in boxes for more detailed race and ethnicity information (e.g., Asian Indian, Chinese, Filipino).

2024 Revisions to Race and Ethnicity Data Standards

In 2024, the Office of Management and Budget released revisions to SPD15. The new minimum standards combine race and ethnicity into a single self-reported question and continue to allow respondents to select multiple categories. These revisions will improve the quality of data on race and ethnicity by giving people better options for identifying themselves. Under the 1997 standards, many people who identify as Hispanic or Latino did not identify with any of the race categories, and there was no distinct category for people who identify as Middle Eastern or North African.

The new 2024 minimum categories are:

- American Indian or Alaska Native
- Asian
- Black or African American
- Hispanic or Latino
- Middle Eastern or North African
- Native Hawaiian or Other Pacific Islander
- White

Data are not yet available under these new minimum standards as implementation plans and guidance are still being developed.

DATA AND TOOLS

Most data sources provide data disaggregated into major race and ethnicity categories as required by the federal government. One challenge when reporting disaggregated data is maintaining data reliability, due to the smaller number of people in each category. Be sure to follow the previously described steps for calculating and assessing reliability, and use aggregation techniques to improve reliability where needed.

Using microdata tools such as MDAT or IPUMS allows for disaggregating data into more detailed race and ethnicity categories with some datasets.

The Census Bureau also produces tables by detailed race and ethnicity with the ACS and decennial census.

ACS Tables

Selected Population Profiles, ACS Table S0201

Selected population profiles are data tables that provide information on broad social, economic, housing, and demographic characteristics for many race, ethnicity, ancestry, and country/region of birth groups (see Figure 12). They are produced annually using the ACS 1-year estimates. These tables have a population threshold of 65,000 or more, meaning they are only available for larger geographic areas and subpopulation groups.

FIGURE 12. SNAPSHOT OF A SELECTED POPULATION PROFILE

S0201 | Selected Population Profile in the United States
American Community Survey | 2023: ACS 1-Year Estimates Selected Population Profiles

Notes · Geos · Topics · Codes · Dataset · Year · Columns · Transpose · Margin of Error · Restore · Excel · CSV · ZIP

Please note that American Community Survey 1-Year estimates are published for geographies with a population of 65,000 or more. For more information, see the guidance for when

Label	United States Black or African American alone Estimate	Margin of Error	Asian Indian alone Estimate	Margin of Error	Native Hawaiian alone or in any combination Estimate	Margin of Error	Chinese, except Taiwanese alone Estimate	Margin of Error
TOTAL NUMBER OF RACES REPORTED								
Total population	40,619,972	±94,692	4,713,643	±62,457	693,679	±22,487	4,228,441	±52,414
One race	100.0%	±0.1	100.0%	±0.1	27.3%	±1.5	100.0%	±0.1
Two races	(X)	(X)	(X)	(X)	35.7%	±1.6	(X)	(X)
Three races	(X)	(X)	(X)	(X)	30.5%	±1.4	(X)	(X)
Four or more races	(X)	(X)	(X)	(X)	6.5%	±0.9	(X)	(X)
SEX AND AGE								
Total population	40,619,972	±94,692	4,713,643	±62,457	693,679	±22,487	4,228,441	±52,414
Male	48.0%	±0.1	52.2%	±0.3	50.0%	±0.9	47.0%	±0.3
Female	52.0%	±0.1	47.8%	±0.3	50.0%	±0.9	53.0%	±0.3
Under 5 years	5.9%	±0.1	5.9%	±0.2	7.7%	±0.6	3.4%	±0.2
5 to 17 years	17.7%	±0.1	16.6%	±0.3	22.8%	±1.0	12.2%	±0.2
18 to 24 years	9.9%	±0.1	8.1%	±0.2	10.2%	±0.6	9.1%	±0.2
25 to 34 years	15.1%	±0.1	17.2%	±0.4	14.4%	±0.8	16.1%	±0.3

Source: U.S. Census Bureau, "Table S0201, Selected Population Profile in the United States," 2023 ACS 1-year estimates.

Selected Population Tables and American Indian and Alaska Native Tables

The Selected Population Tables and the American Indian and Alaska Native Tables provide detailed statistics on social, economic, housing and demographic characteristics for hundreds of race, tribal, Hispanic origin, and ancestry populations for many geographic levels. They are available every five years, with the latest iteration using the 2021 ACS 5-year estimates. These packages contain over 250 iterated tables. Not all population groups are available at all geographies. Table 7 outlines the thresholds that must be met in order for a population group and geography to be available.

CONSIDERATIONS FOR TREND ANALYSIS

Question and data processing changes need to be carefully considered when using disaggregated data in a trend analysis. While such changes might result in a more accurate picture of how people self-identify, they present difficulties in comparing the same group of people over time.

Major changes in how data are collected by race and ethnicity, such as the implementation of the 2024 revision of SPD 15, will result in a break in trend or require the need for linking categories across the two sets of definitions. Details on this process will be forthcoming as SPD 15 is put into practice.

But smaller changes can also impact the usability of disaggregated data for trend analysis. The Census Bureau made changes to the question design,

TABLE 7. STEPS FOR DETERMINING POPULATION GROUPS AND GEOGRAPHIES IN THE SELECTED POPULATION AND AMERICAN INDIAN AND ALASKA NATIVE TABLES

STEP 1. POPULATION GROUPS AVAILABLE	STEP 2. GEOGRAPHIES AVAILABLE
For Selected Population Tables: • The race, Hispanic origin, ancestry, or tribal population group must have a population of at least 7,000 at the national level. **For American Indian and Alaska Native Tables:** • The American Indian or Alaska Native group must have a population of at least 100 at the national level.	• For the table to be released for a particular population group, the desired geographic area must have at least 50 unweighted sample persons.

Source: U.S. Census Bureau, "Race, Ethnicity, Ancestry, and American Indian and Alaska Native Tables," revised October 28, 2024.

data processing, and coding for racial and ethnic categories in the 2020 Census and with the ACS beginning in 2020.

The most notable change from the previous census is a large increase in the multiracial population in the United States, driven by two groups: people who identify as white and at least one additional race and people who identify as a race not listed and at least one additional race. This means that not only has the size of the multiracial population grown, but also the composition of this population has changed relative to 2010.

The Census Bureau has recommended using caution when comparing ACS data with the new race and ethnicity design (from 2020 onward) against data from prior years.

Further Learning

This *KIDS COUNT Data Resource Guide* describes the key data sources used in national-level KIDS COUNT data products, such as the KIDS COUNT Data Center, the *KIDS COUNT Data Book*, and policy reports. The data sources included in this guide are not exhaustive of all sources used by KIDS COUNT.

PRB aims to fully support the KIDS COUNT network across data collection, analysis, and use activities. Please reach out to PRB, specifically Alicia VanOrman (avanorman@prb.org), with any questions about the content in this guide, about using these data sources in your work, and about data needs that can support your work more broadly.

PRB

ABOUT POPULATION REFERENCE BUREAU

Population Reference Bureau (PRB) is a nonpartisan, not-for-profit research organization that uses population data to improve lives across the globe. PRB conducts research and data analysis on population trends and the implications for public policies. PRB is a national data provider for KIDS COUNT, compiling data for the KIDS COUNT Data Center, *KIDS COUNT Data Book*, and KIDS COUNT Policy Reports and Data Snapshots. We are experts in the federal statistical system and apply this expertise to numerous additional data systems.

We also provide technical assistance to support research and data analysis activities, so that organizations have the best population data available to answer pressing questions. PRB can help organizations identify the best data sources and analysis techniques, develop indicator definitions, solve data challenges, and assist with creating custom tabulations, among many other data-related activities.

For more information, visit prb.org.